Road network selection for medium scales using an extended stroke-mesh combination algorithm

Benz, Stefan A ; Weibel, Robert

Abstract: The road network is an essential feature class in topographic maps and databases. Road network selection for smaller scales forms a prerequisite for all other generalization operators and is thus a fundamental operation in the overall process of topographic map and database production. The objective of this paper was to develop an algorithm for automated road network selection from a large-scale (1:10,000) to a medium-scale database (1:50,000). The project was pursued in collaboration with swisstopo, the national mapping agency of Switzerland. Three algorithms (a stroke-based, a mesh-based, and a combined stroke-mesh algorithm) were implemented from the literature and analyzed using swisstopo's large-scale TLM3D spatial database, with requirements set forth by expert cartographers. Initial experiments showed that the combination algorithm performed best, yet still it could not meet all requirements. Therefore, extensions to the basic stroke-mesh algorithm were developed, significantly improving the selection result with real-world, large test databases. Three extensions introduce modifications to the stroke-mesh combination algorithm. Furthermore, two extensions include external feature classes, ensuring accessibility of points of interest and appropriate network density representation in settlement areas, respectively. The results were evaluated by expert cartographers, who concluded that the proposed approach is ready to be deployed in production at swisstopo.

DOI: https://doi.org/10.1080/15230406.2014.928482

Originally published at:
Benz, Stefan A; Weibel, Robert (2014). Road network selection for medium scales using an extended stroke-mesh combination algorithm. Cartography and Geographic Information Science, 41(4):323-339.
DOI: https://doi.org/10.1080/15230406.2014.928482

Benz, S. & Weibel, R. (2014): Road network selection using an extended stroke-mesh combination algorithm. *Cartography and Geographic Information Science*, **41**(4): 323-339.

Road network selection for medium scales using an extended stroke-mesh combination algorithm

Benz, Stefan A. & Weibel, Robert
University of Zurich, Department of Geography
Winterthurerstrasse 190, CH-8057 Zurich (Switzerland)
stefan-benz@hispeed.ch, robert.weibel@geo.uzh.ch

Abstract. The road network is an essential feature class in topographic maps and databases. Road network selection for smaller scales forms a perquisite for all other generalization operators and is thus a fundamental operation in the overall process of topographic map and database production. The objective of this paper was to develop an algorithm for automated road network selection from a large-scale (1:10,000) to a medium-scale database (1:50,000). The project was pursued in collaboration with swisstopo, the national mapping agency of Switzerland. Three algorithms (a stroke-based, a mesh-based, and a combined stroke-mesh algorithm) were implemented from the literature and analyzed using swisstopo's large-scale TLM3D spatial database, with requirements set forth by expert cartographers. Initial experiments showed that the combination algorithm performed best, yet still it could not meet all requirements. Therefore, extensions to the basic stroke-mesh algorithm were developed, significantly improving the selection result with real-world, large test databases. Three extensions introduce modifications to the stroke-mesh combination algorithm. Furthermore, two extensions include external feature classes, ensuring accessibility of points of interest and appropriate network density representation in settlement areas, respectively. The results were evaluated by expert cartographers, who concluded that the proposed approach is ready to be deployed in production at swisstopo.

Keywords: map generalization; road network selection; stroke; mesh; stroke-mesh combination

Introduction

The road network is an essential part of topographic maps and databases (Zhang 2004; Touya 2010), and its generalization is considered to be complex. Roads are connected to each other by road segments (i.e. topologically speaking, the edge between two nodes in the corresponding road network graph), forming a contiguous network. Deletion of road segments without consideration of the entire network will result in loss of connectivity (Chaudhry and Mackaness 2005).

The selection of road segments usually precedes, and is a prerequisite for, other generalization operators (Jiang and Harrie 2004). It aims to reduce the level of detail in the road network by choosing the relevant road segments but also maintaining the main characteristics and structure (Touya 2010), as well as the essential topological, geometrical and semantic properties (Mackaness 1995). In order to achieve this, a hierarchy of roads is necessary. However, this hierarchy is difficult to obtain because it depends on different aspects that are related and coupled, such as geometry, attributes, topology and the implicit geographical context. Furthermore, differences in the road network density distribution in urban and rural areas have to be considered, too (Touya 2010).

There exist several algorithms that perform the selection process automatically. Among these, the stroke-based approach (Thomson and Richardson 1999) and the mesh-based approach (Chen et al. 2009) have been reported to produce promising results. Recently, an algorithm has been introduced that combines both the stroke, and the mesh-based approach in an integrated concept (Li and Zhou 2012). However, this algorithm has not been tested for many test cases yet, and particularly not for large and heterogeneous regions typical of topographic map and database production at national mapping agencies (NMAs) and commercial cartography companies.

This paper derives from a research project (Benz 2013) that was pursued in collaboration with swisstopo, the national mapping agency of Switzerland. The topographic map and database products of swisstopo are derived from the TLM3D (Topographic Landscape Model 3D) database, and swisstopo — like many other NMAs — are seeking to automate their production processes, including road network selection as one example. TLM3D is a highly detailed, dense and large-scale spatial database corresponding to the source scale range of 1:10,000, covering Switzerland and the Principality of Liechtenstein with a high spatial resolution in all three spatial dimensions.

The interest of this project was on the transition from a large-scale source database to a medium target scale of 1:50,000. TLM3D was used as an example source database, but we believe our results to be adaptable to other databases of other NMAs as well. An initial analysis was conducted of how well three different algorithms for road network selection (the stroke-based, mesh-based, and the combined stroke-mesh algorithm) perform for this scale transition. The selection process and the algorithms were implemented using the constraint-based generalization approach as proposed by several researchers (Beard 1991; Weibel and Dutton 1998; Harrie and Weibel 2007). First, constraints set forth by expert cartographers were taken into account. Second, the results were evaluated using the defined constraints. The aim set by swisstopo is to select 70 % of the features from the source database for the target scale.

A first round of experiments, documented in detail in (Benz 2013), showed that the integrated stroke-mesh algorithm by Li and Zhou (2012) clearly produces the best results of the road network selection algorithms published to date. As will be explained in the following section, this algorithm represents a combination of two important streams of algorithms, the stroke-based and the mesh-based principle, in an attempt to capitalize on their respective strengths. Thus, invariably, results generated are superior to those obtained from either the stroke- or mesh-based algorithms alone, as documented by Li and Zhou (2012) for real test cases of considerable size. Medium scales, such as the target scale of 1:50,000, are still rather detailed and thus road network selection should handle the details of the network carefully, which turns out to be a particular strength of the stroke-mesh algorithm. Additionally, the algorithm can maintain the general connectivity of the network (i.e. no new dead-end roads arise where there are actually none), which was one of the main constraints defined by the swisstopo experts, and which is a general requirement in map generalization.

Despite these advantages, our preliminary experiments revealed that several difficulties remained, and not all of the cartographic requirements could be fulfilled. Therefore, an in-depth analysis was carried out as to how additional constraints and extensions could improve the results. These constraints and extensions form the focus of this paper.

The main contribution of this paper thus consists of five extensions that address deficiencies of the basic version of the integrated stroke-mesh algorithm by Li and Zhou (2012) and lead to a significant improvement of the results. Two of these extensions apply directly to the inner workings of the basic stroke-mesh algorithm. A third extension provides an approach to ensure that roundabouts are selected or omitted correctly (i.e. selected or omitted as complete objects). Additionally, two extensions are introduced that allow incorporating external feature classes. External feature classes are seldom used in the automated selection process of road networks. However, such external classes contain additional information that can guide and improve the selection. The first class used contains a set of POIs (points of interests) and the second class consists of a set of polygons that define settlement areas. The POIs are used to ensure the accessibility to important infrastructure objects in the selected road network whereas the settlement areas are used to guarantee that a cartographically adequate road network density distribution can be maintained in the result. For situations where a settlement feature class is not available we also show how proxies of settlement areas can be generated from the road network.

The results of the extended stroke-mesh combination algorithm have been thoroughly evaluated in four large test areas of largely differing and heterogeneous road network character, both by quantitative analysis and by a detailed assessment by swisstopo experts.

The experts' evaluation was highly positive, leading to the conclusion that the proposed approach is ready to be deployed in production at swisstopo.

In the next section, we will discuss the state of the art that also provides a short overview of the stroke-mesh combination algorithm. The third section then introduces our extensions to the basic stroke-mesh combination algorithm, one by one. The fourth section presents the evaluation of the final results by swisstopo experts. The paper ends with conclusions and an outlook on future research.

State of the art

Related work

Liu, Zhan, and Ai (2010) subdivide automatic road network selection into three groups: semantics-based selection, graph-based selection and stroke-based selection.

Semantics-based selection makes use of attributes (e.g. road class). Roads are ordered according to their relative importance of attributes and the selection is based on this order. Owing to their simplicity, they are often used in practice. However, they are clearly insufficient due to the neglect of geometrical and topological constraints.

The graph-based methods treat road networks as connected graphs and use pattern detection algorithms (Yang, Luan, and Li 2011) or concepts such as shortest/best path or minimum-spanning-trees, which serve as the basis for the selection (Mackaness and Beard 1993; Mackaness 1995). A special group of algorithms uses the dual graph approach at the topological level, where nodes represent roads and edges represent intersections of roads, respectively (Porta, Crucitti, and Latora 2006). Dual graphs are often used to compute centrality measures, such as degree, closeness or betweenness as a measure of importance for roads which then serve as the basis for the actual selection task (Jiang and Claramunt 2004). With their focus on graph-theoretic principles that involve the entire road network, such as centrality measures, graph-based algorithms tend to favor the main connectivity structures of the network, making them particularly suitable for selection at smaller scales, as shown by a related study (Weiss 2013).

Stroke-based selection is based on the principle of ‘good continuation’ and takes into account functional importance and perceptual significance. A stroke is a chain of road segments with continuous curvature, that is, with ‘good continuation’ (Thomson and Richardson 1999; Thomson and Brooks 2000). Strokes can be used for pattern detection (Heinzle, Anders, and Sester 2005; Thom 2005; Yang, Luan, and Li 2011), for topological analysis (Touya 2010) or for the hierarchical modeling of roads (Tomko, Winter, and Claramunt 2008; Jiang 2009). Most often, however, strokes are used for the selection of road networks (Thomson and Richardson 1999; Edwardes and Mackaness 2000; Thomson and Brooks 2000; Chaudhry and Mackaness 2005; Heinzle, Anders, and Sester 2005; Liu, Zhan, and Ai 2010; Yang, Luan, and Li 2011; Li and Zhou 2012). The strokes are ordered according to some predefined rules (e.g. length or attributes of strokes), and the selection is conducted selecting strokes in descending order. The general idea is that long strokes with attributes of higher order form roads with a high functional importance that should be selected for the smaller target scale, whereas short strokes with attributes of lower order may be omitted.

There exist additional algorithms that cannot exactly be matched to one of the groups defined by Liu, Zhan and Ai (2010). For instance, Morisset and Ruas (1997) measured the importance of roads based on how often they are likely to be used by means of an agent-based simulation.

Another concept used is the mesh-based approach proposed by Chen et al. (2009), which considers the areal properties in a road network. The concept is quite similar to the method proposed by Edwardes and Mackaness (2000) but has been elaborated in more detail. A mesh is a closed region that is bounded by several road segments. Figure 1 provides a hypothetical road network with five meshes. The selection is based on the identification of meshes with a high mesh density, which is defined as the ratio of the perimeter and the area of the mesh. In Figure 1, the mesh with the highest density (mesh number 3) is treated first. Its bounding segments are ordered according to their relative importance and the least important segment (red) is eliminated first. The remaining segments are merged with the adjacent mesh (mesh number 2), thus forming a new mesh with a lower mesh-density. This process is repeated until all meshes have a mesh-density smaller than a predefined threshold or the desired number of segments has been omitted.

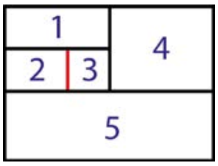

Figure 1. A hypothetical road network with five meshes.

Integrated stroke-mesh combination algorithm

Li and Zhou (2012) have proposed an algorithm that combines both the mesh-based and the stroke-based approach in an integrated concept, thus seeking to capitalize on the strengths of both approaches. We will present a short summary of the most important concepts of the combined algorithm by Li and Zhou (2012), as the understanding of this algorithm is necessary to understand the improvements and new concepts to extend their previous work. A more detailed description, though, is given by Li and Zhou (2012).

In order to understand the principle of the integrated approach, one has to differentiate between *areal* and *linear* segments, respectively. An areal segment is a segment that forms the boundary of one or two meshes, as is each segment in Figure 1 (together they form a so-called areal pattern). A linear segment is a segment that does not belong to any side of a mesh, such as a single stroke or a dead-end. Figure 2 provides an example of a road network consisting of linear segments only (jointly forming a linear pattern).

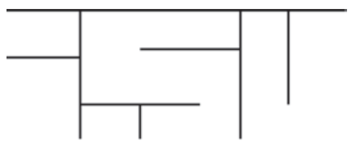

Figure 2. A hypothetical road network with linear segments only.

Li and Zhou (2012) have shown that the stroke-based approach performs best for linear patterns and the mesh-based approach for areal patterns, respectively. Therefore, their integrated approach handles linear segments with the stroke-based approach and areal segments with the mesh-based approach.

The basic idea is to build up a hierarchical structure of a road network. An areal hierarchy consisting of only areal segments is constructed using the mesh-based approach. Thus, the meshes are dissolved and merged with neighboring meshes until there is only one large mesh left (the outer mesh). Every step is recorded and stored in a tree data structure.

A linear hierarchy that consists of linear segments only is then constructed using the stroke-based approach. The segments are concatenated into strokes, which then again are ranked according to their topological depth and placed into another tree data structure. Figure 3 shows a linear pattern. For a purely linear pattern, the starting stroke is the longest stroke (stroke 1). Hence, it builds the root node of the tree. Stroke 3 is connected to stroke 1 and thus a child node of stroke 1. Stroke 2 and stroke 5 both connect to stroke 3 and thus are both children-nodes of stroke 3. Finally, stroke 4 connects to stroke 5 and thus is a child-node of stroke 5.

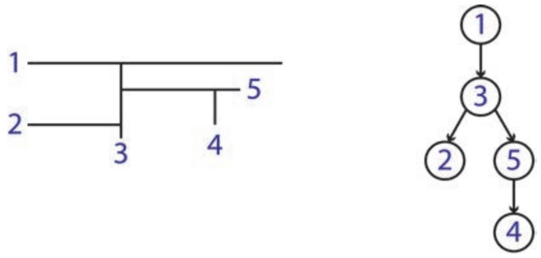

Figure 3. Linear pattern of strokes and the corresponding tree data structure.

Mostly, road networks do not form purely linear or areal patterns, respectively, but hybrid patterns with both linear and areal segments. Therefore, linear and areal hierarchies need to be connected. This connection is done using so called *bridges*. Bridges appear at locations where linear and areal patterns are connected to each other. Such a situation is depicted in Figure 4. The linear pattern is connected to an areal pattern at the red point. The blue stroke builds a bridge as it ensures that the linear and areal patterns are connected.

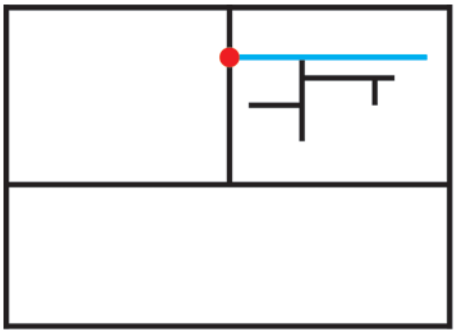

Figure 4. Linear pattern of strokes and the corresponding tree data structure.

The actual selection is then performed by traversing these hierarchies. Basically, in the first step, the areal hierarchies are traversed until the desired number of areal segments has been selected (or omitted, respectively). This can be done by using a predefined threshold for the mesh density, meaning that each mesh must not exceed a certain density or by simply using a predefined number of segments that must be selected.

In the second step, the linear hierarchies are traversed. If such a hierarchy is connected to an areal segment that was *retained* in the first step, the stroke that is connected to this areal segment (the bridge) builds the root node and all other strokes form children- and grandchildren-nodes as depicted in Figure 3. The traversal of the tree starts at the root. If the length of the currently traversed stroke is longer than a specific threshold, the stroke is selected. The traversal then continues according to the same principle from this parent stroke to its children strokes, until the currently traversed stroke has no more child strokes, or its length is shorter than a predefined threshold.

We will now proceed to the detailed analysis of the stroke-mesh algorithm, with an emphasis on those steps that violate key cartographic requirements, thus introducing extensions that improve the original algorithm by Li and Zhou (2012).

Extensions of the integrated approach

Cartographic constraints

Before we start with the description of our extensions, we briefly need to define the cartographic requirements that the results of the road network selection process should meet, in order to assess what the deficiencies of the original stroke-mesh combination algorithm are, and how they could be improved. In consultation with the expert cartographers at swisstopo, cartographic requirements were defined for the transition to the target scale of 1:50,000, leading to a set of hard and soft constraints, respectively. Hard constraints are mandatory and must be fulfilled by the result. Soft constraints, on the other hand, are constraints that the result should fulfill as far as possible. Evaluation of hard constraints leads to binary decisions, while in soft constraints the degree of fulfillment is assessed.

Hard constraints:

- (H1) The fraction of road segments retained at the target scale should be 70 % of the segments at the source scale.
- (H2) Road segments that contain certain TLM3D attributes (highways, connection-roads, long tunnels etc.) must be retained.
- (H3) Roads must not be disconnected locally. Therefore, no new dead-ends must be generated, except if they connect to important infrastructure objects.
- (H4) A segment or a set of segments must not be isolated from the rest of the network, meaning they need to remain connected to the network, unless they were already isolated in the source database.

(H5) Roundabouts must not be collapsed; they must either be removed or selected as a whole.

(H6) Important infrastructure objects and larger areas need to remain accessible in the selected road network.

It is debatable whether H1 should be considered as a hard constraint because one might argue that a result in the range between 68 % and 72 % could be equally acceptable. In our case, however, after discussions with swisstopo experts, we defined it as hard because the algorithm should come as close as possible to the 70 % mark.

Soft constraints:

(S1) Roads that a map user would subjectively interpret as important for the overall structure of the road network should be selected.

(S2) The general structure and character of the road network should be maintained.

Note that the above requirements, although specified by swisstopo, represent common cartographic practice and are easily adapted to the particular situation at other mapping agencies. For instance, the threshold of 70 % of retained road segments in constraint H1 is rather high, and thus might be chosen to be lower by other mapping agencies. Similarly, the list of priority road features in H2 could be adapted to the priority features at other organizations. The algorithmic extensions proposed in this paper can thus be assumed to be generalizable with little effort.

The implementation of our extended algorithm is configured taking into account these constraints. For instance, segments that contain important attributes (H2) are not deleted by the algorithm. H1 is achieved iteratively. Specifically, the algorithm can be configured such that it deletes a specific amount or percentage of road segments in the mesh elimination process. Additional segments are deleted in the second part of the combination algorithm. Conversely, some roads are selected again when using our extensions described in the following sections. Thus, because the algorithm cannot know beforehand how many segments are eliminated and selected in the second part of the combination algorithm and in our extensions, H1 is achieved iteratively by varying the configuration until H1 is fulfilled.

Bottom-Up approach for the traversal of linear hierarchies

Li and Zhou (2012) use a *top-down* approach in the traversal of the tree hierarchy for linear hierarchies. The traversal starts at the root node. If the corresponding stroke is longer than a predefined threshold, it is selected; otherwise the traversal stops. In some cases, however, this approach is problematic. Such a case is sketched in Figure 5. At the top, a linear hierarchy with five strokes is depicted. The selection process now starts at the root node. If we assume that stroke 1 is longer than the length threshold, it is selected and the traversal continues with stroke 3, colored in red. If this stroke is shorter than the threshold, the traversal stops and it is not selected. The result can be seen on the bottom of Figure 5. Only stroke 1 is selected. The entire rest of the linear hierarchy has been omitted.

The problem is that the strokes at the bottom of the tree hierarchy are not considered in the top-down traversal order. Let us assume that stroke 4, colored in green, is longer than the length threshold, contains segments with attributes that make a selection mandatory (constraint H2), or might subjectively be interpreted as important by a map user (constraint S2). Using the top-down approach, this stroke cannot possibly be selected.

Multiple solutions are possible order to solve this problem. The simplest approach would be to select stroke 4 after the traversal of the tree hierarchy. However, this leads to a connectivity problem. Stroke 4 would be disconnected from the rest of the tree hierarchy. This is a violation of the connectivity constraint H3 and H4, defined by the swisstopo experts.

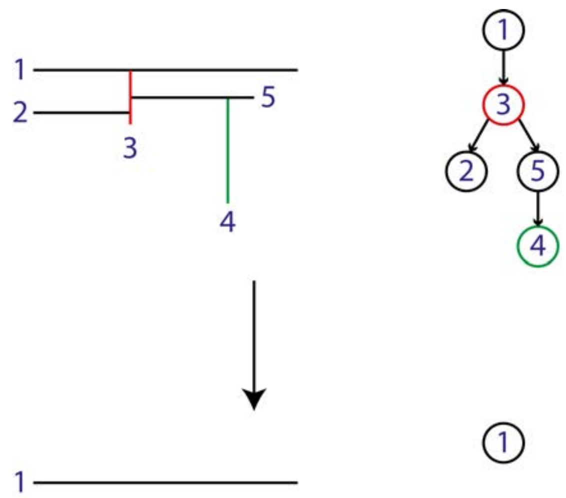

Figure 5. Problematic aspect of the top-down approach for linear tree hierarchies.

Therefore, we propose using a *bottom-up* approach where the traversal starts at the leaf nodes of the tree. If a leaf node (i.e. the corresponding stroke) is longer than the length threshold or contains attributes that make a selection mandatory, the stroke itself *and* all its parent and grandparent strokes are retained, even if they are shorter than the length threshold. The traversal then continues according to the same principle with all the parent and grandparent-nodes of the leaf nodes.

This approach ensures that all important strokes are selected, even if they appear at the bottom of the tree hierarchy, and it guarantees that these strokes are connected to the network through their parent and grandparent strokes, thus fulfilling constraint H3 and H4. If this approach is applied to the situation in Figure 5, it leads to the result shown in Figure 6. The traversal starts at the leaf nodes (strokes 2 and 4). Whether the traversal starts with stroke 2 or stroke 4 does not matter. The result will be the same. If we assume that stroke 4 is longer than the threshold or needs to be selected because of its attributes, the stroke itself, as well as all its parent- and grandparent strokes (strokes 5, 3 and 1) are selected. Stroke 2 is supposed to be shorter than the length threshold, therefore it is not selected. The traversal then continues with the strokes 5 and 3. Both are already selected, so they can be skipped. The same applies to the root node (stroke 1). The resulting situation is depicted at the bottom in Figure 6. The important stroke 4 is selected and remains connected to the rest of the network.

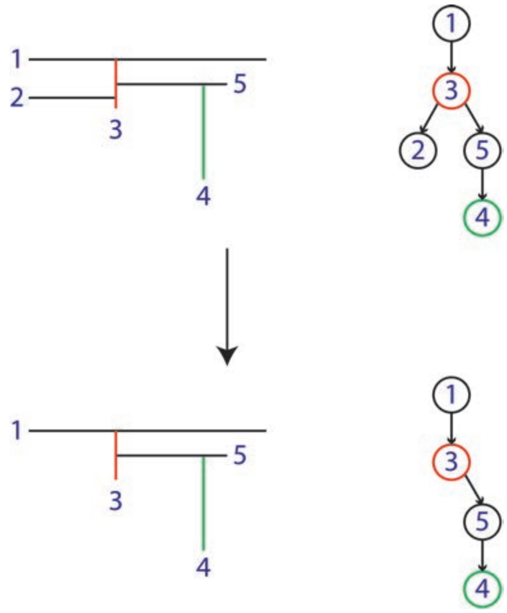

Figure 6. Solution using the bottom-up approach for linear tree hierarchies.

Figure 7 shows snippets from two of our test areas. Both examples contain linear patterns (colored in red) where it makes a difference whether the top-down or bottom-up approach is applied. Figure 7(a) depicts a linear pattern in a mountainous area. The hierarchy is connected to a mesh at the southern end. The root node of the hierarchy is rather short. Using the top-down approach, it is not selected and neither is the rest of the pattern because the traversal already stops at the (too short) root node. The strokes at the bottom of the hierarchy, however, are much longer than the length threshold and need to be selected because they can be interpreted as important roads making a large area accessible (constraint H6) or as subjectively important roads mainly due to their length (constraint S1). A simpler linear pattern with two strokes in a rural area is depicted in Figure 7(b). Once again, the root node is shorter than the length threshold. However, the single child node is longer and has to be selected. The bottom-up approach guarantees the selection of the long strokes and its connection to the rest of the network, while the top-down approach obviously does not.

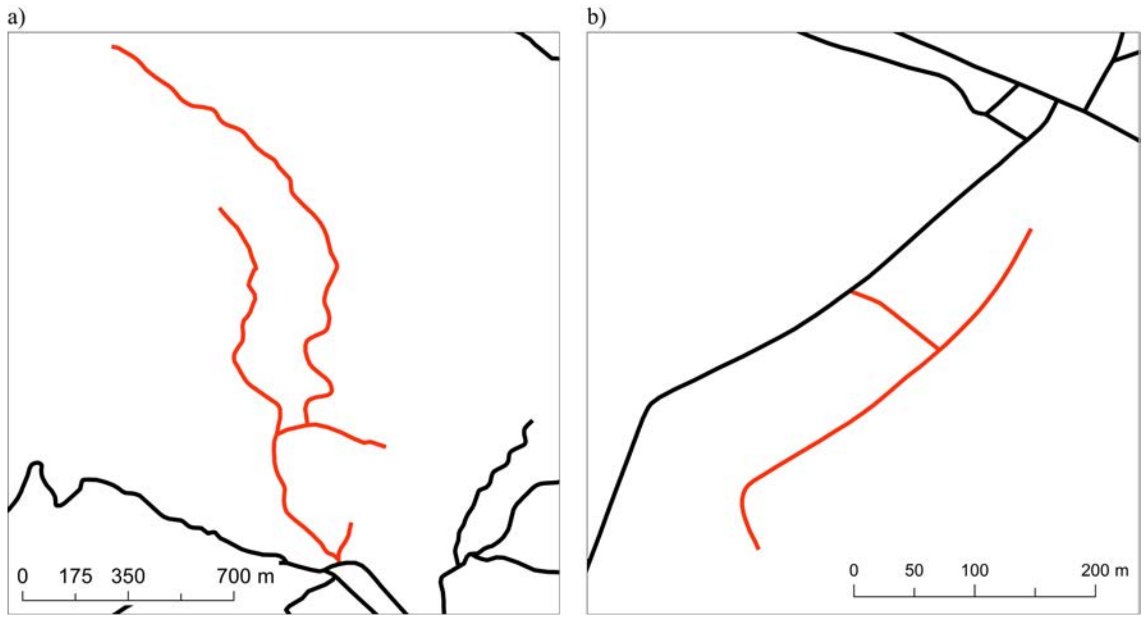

Figure 7. Two road network snippets illustrating the resulting difference of the top-down- and bottom-up approach (© swisstopo).

Linear patterns connected to dense meshes

The integrated approach by Li and Zhou (2012) does not handle linear patterns that are connected to a mesh that was dissolved in the first step of the integrated approach. A linear pattern is only considered *if* it is connected to a mesh that was retained in the first step (i.e. in the mesh generalization phase). While this approach might be applicable in urban areas, our experiments showed that overall it tends to generate insufficient results, especially in suburban, rural and mountainous areas.

Figure 8 shows two road network snippets of the result of the integrated approach for two of our test areas. The omitted segments are colored in red. In Figure 8(a), it can be seen that there is e mesh that was dissolved (marked by a blue ellipse) and one very long singular stroke that is connected to that mesh. Because the integrated approach does not handle such linear patterns, it is omitted as well, even though it should be selected for the 1:50,000 scale mainly because it could be interpreted as an important road because of its length (constraint S1) and because it makes a larger area accessible (constraint H6). A more complex situation can be observed in Figure 8(b). There are two dissolved meshes (again marked by blue ellipses) and several linear patterns connected to those meshes. The linear patterns extending beyond the dense meshes are not selected due to the aforementioned problem, even though they make a large area accessible.

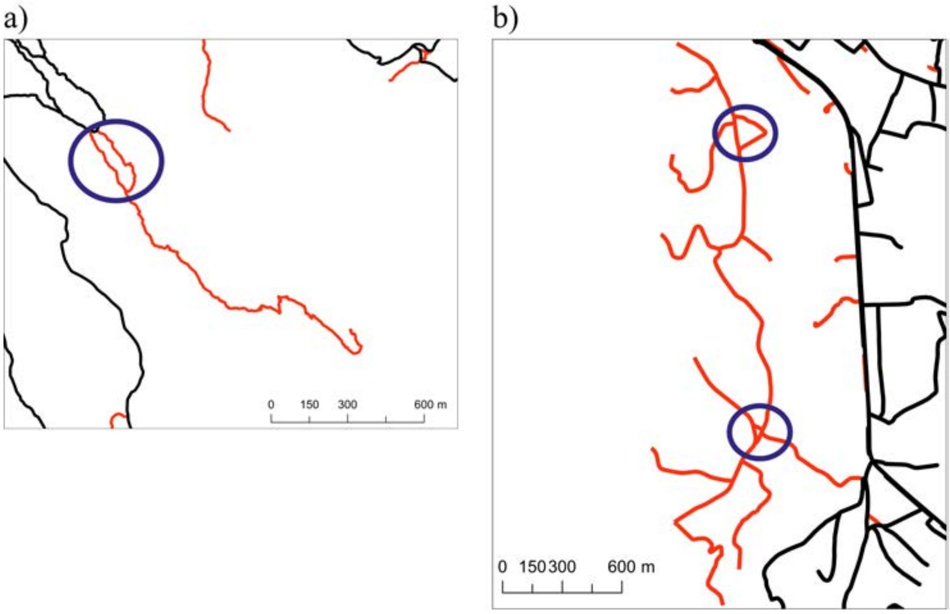

Figure 8. Two road network snippets illustrating the problem with linear patterns at dense meshes (© swisstopo).

In order to resolve this problem, multiple solutions are imaginable. One such solution would be to add the long strokes after running the basic stroke-mesh algorithm. However, this would once again impose a connectivity problem because the strokes would not necessarily be connected to the rest of the network, thus violating constraints H3 and H4. We therefore implemented a stroke-reconnection algorithm that uses shortest path concepts. Specifically, if there is a stroke within a linear pattern that is longer than the threshold or contains attributes that make a selection mandatory, but not connected to a mesh that was retained in the first step, the shortest path (Hart, Nilsson and Raphael 1968) in the source database (i.e. the complete TLM3D road network) is computed between this stroke and the nearest *areal* segment selected in the first, mesh-pruning step of the integrated approach. This path can then be used to select the stroke itself but also the segments of the shortest path. This ensures that the stroke is connected to the rest of the road network, thus fulfilling the constraints H3 and

H4. Figure 9 shows the result of this approach. The long stroke in Figure 9(a) is now selected and connected to the rest of the network, while the mesh is still partly dissolved. The same principle can be seen in Figure 9(b). The longer strokes within the linear patterns are selected and connected to the rest of the road network in the East, while the two meshes are partly dissolved.

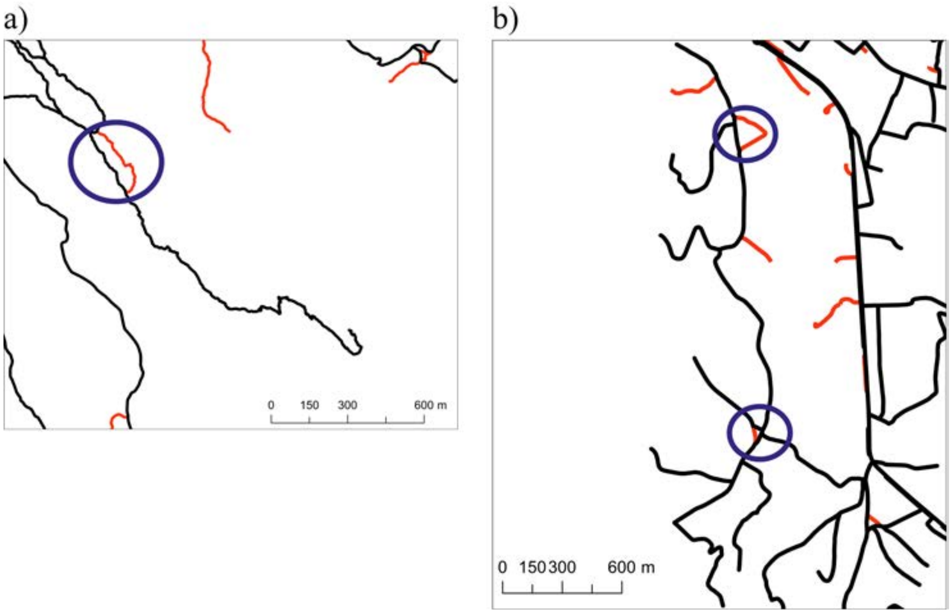

Figure 9. The results using the stroke-reconnection algorithm (© swisstopo).

For the 1:50,000 target scale, the simple shortest path proved to be applicable since mostly there is only one path that connects the stroke to the road network anyway, which is due to the fact that the selection percentage H1 (70 %) is relatively high, and thus not that many segments are omitted. However, for smaller target scales with a much smaller selection percentage, there might be a larger set of possible paths, and it would then be necessary to search for the most important path from a structural and topological view.

Roundabout correction

Road networks, particularly in European countries, usually contain roundabouts. Since they are closed, circular shapes, they form dense meshes and the corresponding segments are therefore prone to elimination using the integrated or purely mesh-based approach by Li and

Zhou (2012) and Chen et al. (2009), respectively. As a consequence, inadequate results occur as depicted in Figure 10. One segment of the roundabout is omitted (colored in red), while the others are selected. In swisstopo maps at the scale of 1:50,000 (but also in maps of the same scale of other NMAs), roundabouts are not collapsed but shown as complete objects (constraint H5). Hence, a desirable result either selects or omits the entire roundabout. Thus, if only parts of a roundabout were selected using the integrated approach, all the other segments of the roundabout should be selected afterwards as well. In order to do that, however, the roundabouts first have to be identified and extracted.

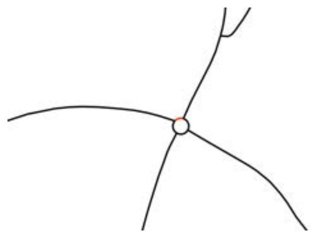

Figure 10. Insufficient result of a roundabout mesh elimination (© swisstopo).

Multiple algorithms have been reported in the literature to detect roundabouts or complex road junctions. Mackaness and Mackechnie (1999) use graph-theoretic principles where a relatively dense collection of vertices, each with degree three or more, are interpreted as a complex junction. Yang et al. (2011) use density-based clustering, that is, an extended DBSCAN algorithm (Ester at al. 1996), while Touya (2010) uses a measure of polygon compactness that is applied on all the small meshes to detect complex junctions.

Unfortunately, it has not been clearly documented how well the above algorithms perform for the detection of roundabouts. Furthermore, we needed an algorithm that strictly finds roundabouts and no other complex junctions, with minimal overhead, re-using structures that had already been computed in preceding steps, most notably the strokes. Therefore, we developed our own approach (jointly with Row Weiss; Weiss 2013). Our algorithm builds directly on the generation of strokes and thus adds little extra effort, both in terms of code development as well as computational effort. It is able to detect roundabouts, excluding other types of complex junctions. In Figure 11, a typical roundabout is shown, where the strokes are colored differently. It can be seen that the roundabout results in a single stroke that forms a loop. This is due to the fact that the segments of the roundabout usually have a small deflection angle and therefore are concatenated into a single and circular stroke. Thus, in order to find all the roundabouts, the algorithm looks for loops inside the strokes. Additionally, a length constraint parameter is necessary in order not to detect strokes that build loops but are not roundabouts. Empirical tests established that 90 m is an optimal threshold because the circumference of roundabouts in Switzerland usually does not exceed that value. For other test areas, however, this threshold might differ. Since it is based directly on the formation of strokes, the approach is extremely fast and was able to robustly detect all roundabouts in the four test areas (a total of 92 roundabouts).

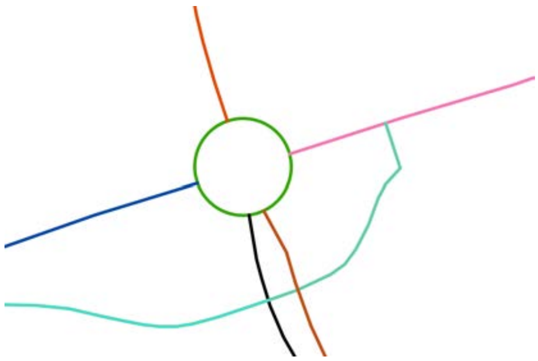

Figure 11. Roundabout and generated strokes (© swisstopo).

Accessibility of points of interests (POIs)

Most approaches for an automated selection of road networks rely on the road network alone. To the authors' knowledge, external feature classes are seldom used. However, the quality of the result can be expected to improve if other feature classes were incorporated because they contain additional information that the road network alone cannot deliver.

One such feature class is a layer of points of interests (POIs) that was provided by swisstopo (and exists similarly at other NMAs). POIs are entities modelled as points having a certain importance. Examples are tramway stations, restaurants in touristic areas, sports arenas and ballparks, hospitals, etc. Depending on the target scale, POIs need to be accessible even in the generalized network (constraint H6). This is certainly the case for the 1:50,000 target scale, which is still rather detailed in its depiction of infrastructures.

In order to develop an algorithm that ensures the accessibility of POIs, one has first to define what 'accessibility of a POI' exactly means. We define a POI as accessible if the segment with the smallest Euclidian distance in the source database (the complete TLM3D in our case) is selected and connected to the main network. In Figure 12, two road network snippets with the result of the integrated approach without considering POIs are depicted. The omitted segments are colored in red, and the POIs are shown by blue point markers. As can be seen, the POIs are not accessible anymore.

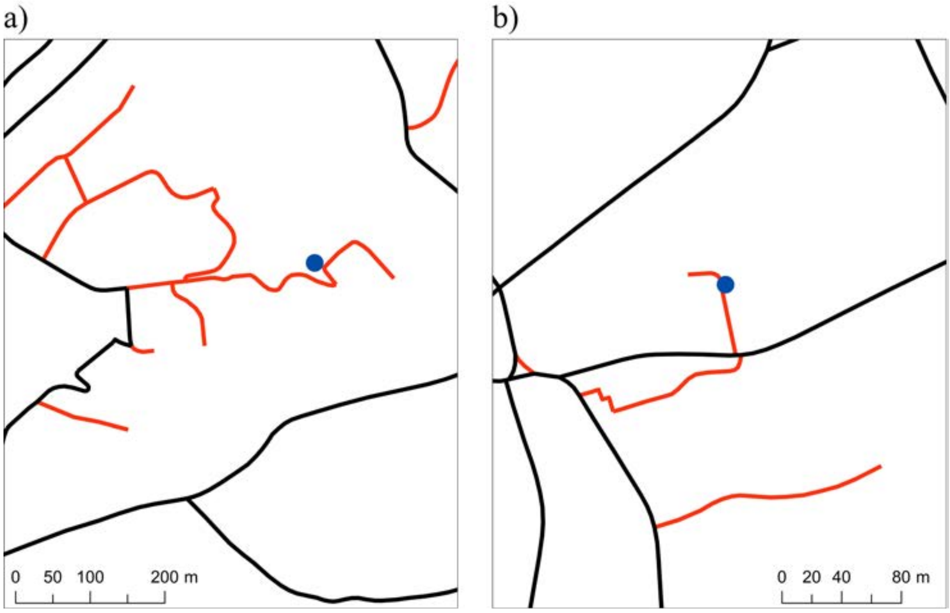

Figure 12. Two road network snippets showing inaccessible POIs (© swisstopo).

The following algorithm describes a mechanism to keep the POIs accessible in the generalized network. It is similar to the approach used by Richardson and Thomson (1996) and Touya (2010). With the difference being that they calculate the shortest paths between all the POIs, while our approach checks for each POI individually how it can re-connect to the main network if the connection was broken.

After using the basic integrated approach, it can be checked whether the segment that has the smallest Euclidian distance to the POI in the complete source database was selected. If so, nothing needs to be done, the POI is either accessible or it was not accessible in the source database in the first place. If the nearest segment was not selected, however, the POI might not be accessible. In this case, the shortest path in the source database is computed from the segment in the source database nearest to the POI and the segment in the *generalized* database nearest to the POI. This path can then be used to ensure that the POI remains accessible. In this case, swisstopo cartographers even allow generating new dead-end roads because they lead to important infrastructure objects (constraint H3).

The two road network snippets in Figure 13 show the result of this approach. For the situation in Figure 13(a) the path to ensure the accessibility contains more than one segment, whereas for the POI in Figure 13(b) only the nearest segment had to be selected afterwards. Both types of cases appear regularly in our four test areas. The first case is particularly frequent in mountainous areas, where access to remote infrastructures and POIs is an important requirement.

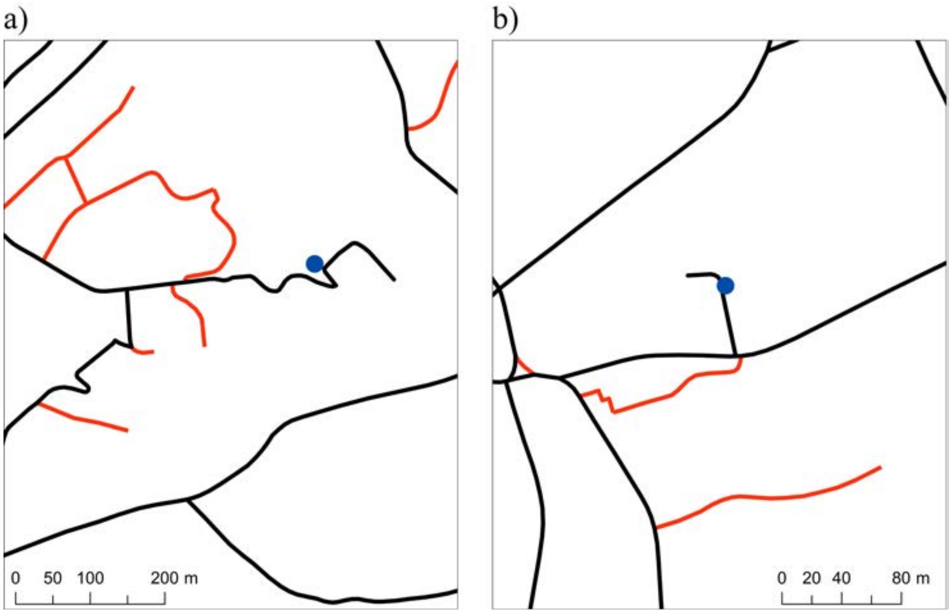

Figure 13. The result of the POI accessibility algorithm (© swisstopo).

One could argue that the shortest path might not be the most important path regarding the usage. However, if the target scale is relatively large (in this case 1:50,000), the results showed that the shortest path is applicable, because it usually does not contain many segments. Additionally, in such situations there are not a large number of other paths that would connect the segment nearest to the POI to the rest of the network. For smaller target scales where more segments are omitted, however, the shortest path might not be the best solution, since there could be other paths that are more important from a topological and structural point of view.

Road network density in settlement areas

Main idea

The first step in the integrated approach of Li and Zhou (2012) thins out dense meshes by eliminating areal segments and merging adjacent meshes. Usually, dense meshes are located in urban settlement areas. For instance, in Figure 14(a) the settlement area in the Northeast has a high number of dense meshes. The consequence is that settlement areas are generalized to a greater extent than the surrounding rural areas and the density of the meshes is more or less evenly distributed in the result (Figure 14(b)), essentially leading to a density equalization, which is cartographically unsatisfactory and violates constraint S2. The settlement areas have been thinned out to such an extent that they are no longer recognizable as such. However, cartographic practice requires the main characteristics and structure of the original network to be retained. Preliminary feedback by swisstopo experts confirmed that assumption.

Figure 14. Source database (a) and result without considering settlement areas (b) (© swisstopo).

In the approach by Li and Zhou (2012) the global parameter based upon which a mesh is eliminated and merged with an adjacent mesh is the mesh density. A natural line of thought, therefore, is to subtract a constant (denoted here as mesh density factor) from a mesh's density if it is in a settlement area. This ensures that an areal segment of a mesh in a settlement area is less likely to be omitted.

One way to decide whether a mesh is in a settlement area is to use an additional feature class that contains the settlement areas modelled as polygons. A mesh is considered to be in a settlement area if one or more of its bounding areal segments are contained within a settlement polygon. Figure 15(a) shows the settlement area layer used for generating the result depicted in Figure 15(b), applying a mesh density factor in settlement areas. The settlement areas now remain clearly recognizable in the road network. The main structure is maintained. As a consequence, the surrounding rural areas are pruned to a greater extent. However, this cannot be avoided if the requirement is to omit a fixed amount (30 % in our case) of the segments from the source database (constraint H1). Thus, it is necessary to find a proper balance of segment elimination in urban and rural areas, respectively.

The question, then, arises how to choose the appropriate value for the mesh density factor. A model that provides a quantitative measure to achieve a proper balance (i.e. to find an optimal mesh density factor) is based on the ratio of the number of segments in urban and rural areas. Figure 16 shows, for the example of one test area with mixed urban and rural landscape, how this ratio depends on different values for the mesh density factor. The red line represents the ratio in the source database, whereas the blue points show the ratio in the generalized result, depending on the chosen mesh density factor. Our model now chooses the optimal mesh density factor such that the ratio from the source database is retained in the generalized result. As can be seen in Figure 16, the linear regression line fits the empirical data almost perfectly ($R^2 = 0.99$).

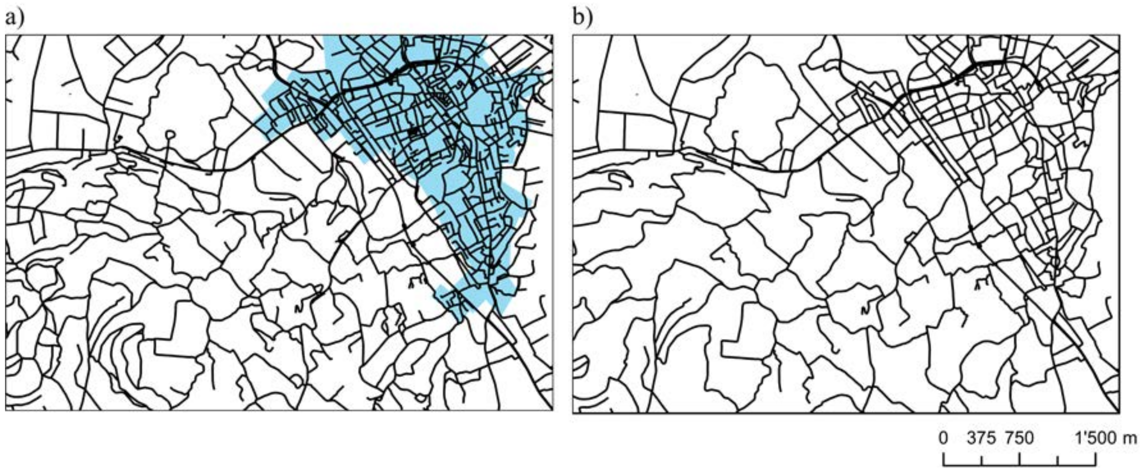

Figure 15. Settlement areas layer (a) used to generate a proper balance of segment elimination in urban and rural areas (b) (© swisstopo).

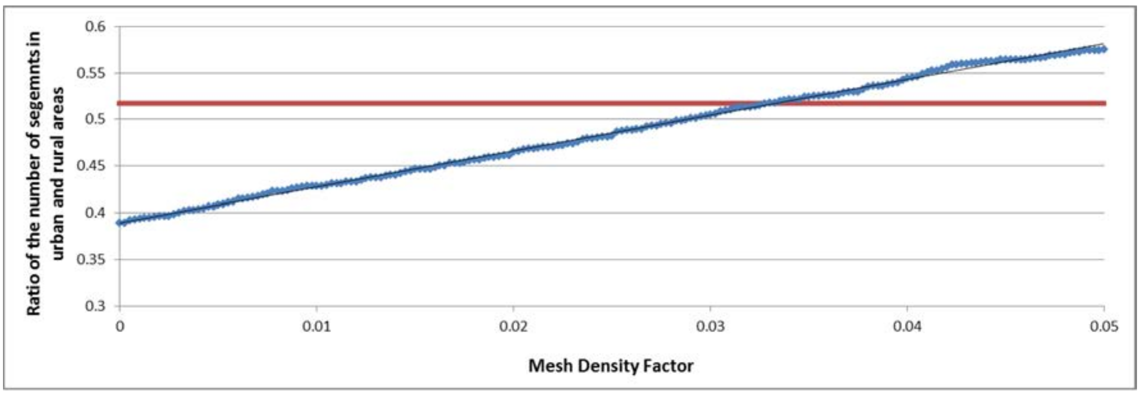

Figure 16. Ratio of the number of segments in urban and rural areas. The red line represents the ratio of the input database, whereas the blue points show the ratio in the generalized result based on the chosen mesh density factor.

Calculating settlement areas from the road network

For the case where no settlement layer is available, a density algorithm is needed that extracts settlement areas (i.e. segments in settlement areas) from the road network itself. Walter (2008) introduced a method to derive raster-based clusters of different degrees of urbanity. This method might be applicable but has the disadvantage of being based on the space-primary principle of the raster data structure. Since we are dealing with vector data structures (like in most cartographic databases), we developed an algorithm that follows the object-primary principle and has the advantage of using a density model that allows a proper configuration of the necessary parameters. Our algorithm is closely related to the KDE (kernel density estimation) approach.

First, the centroids are calculated for each segment, i.e. for each edge in the corresponding graph. Afterwards, for each centroid the number of centroids within a certain radius r is counted. If that count exceeds a predefined threshold g, the centroid (i.e. the segment) is considered to be in a settlement area. In addition, experimental testing has shown that centroids corresponding to highways or segments from roundabouts should not be included in the count in order to achieve better results. Usually, highway segments are relatively long and therefore a centroid is not a really representative model of a road segment, whereas roundabouts form small clusters even in rural areas, because they are made up mostly of four or even more segments, and thus could be mistaken for small settlement areas.

The proposed algorithm has two input parameters (the radius r and the threshold value g) and the question arises how to choose those parameter values, as they are crucial in order to obtain a good result. In order to find optimal parameters, a similarity measure (Li and Zhou 2012) as defined in Equation (1) was used,

$$similarity = \frac{A \cap B}{A + B - A \cap B} \quad (1)$$

where A is the number of segments which according to the algorithm are within a settlement area, and B is the number of segments that are actually contained within a settlement area (the settlement area layer was used as benchmark). Finally, $A \cap B$ is the number of segments which are contained in settlement areas according to the algorithm *and* the benchmark layer. A similarity value of 1 means a perfect solution, whereas a similarity value of 0 corresponds to the worst possible solution.

In order to evaluate optimal parameters, the similarity value was calculated with different values of r and g. Specifically, r varied in the range of 200 m and 700 m with a step size of 10 m, whereas g varied between 30 and 150 and a step size of 2. Figure 17 visualizes the calculated similarity values for one of our four test areas (shown in Figure 18). The pattern is consistent with the pattern for the other three test areas. It shows high similarity values in a strip that extends from pairs of small input parameters to pairs of large input parameters. Low similarity values are generated when using a high value for g and a low value for r, respectively. The maximum similarity of 0.74 is achieved with a radius r of 420 m and a threshold g of 64.

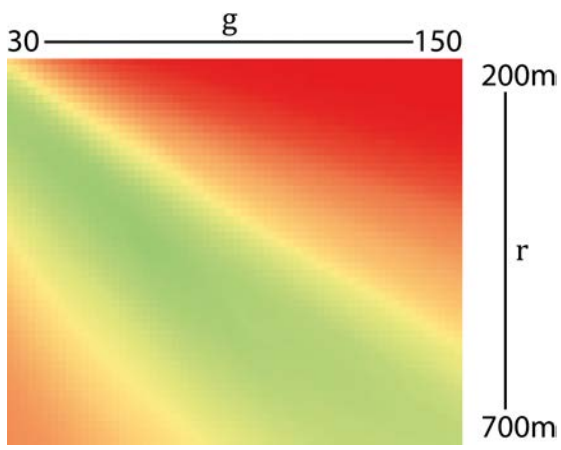

Figure 17. Visualization of the similarity values. Green corresponds to high similarities whereas red symbolizes low similarities.

The result of the density algorithm with this optimal parameter setting for one of our test areas is depicted in Figure 18.

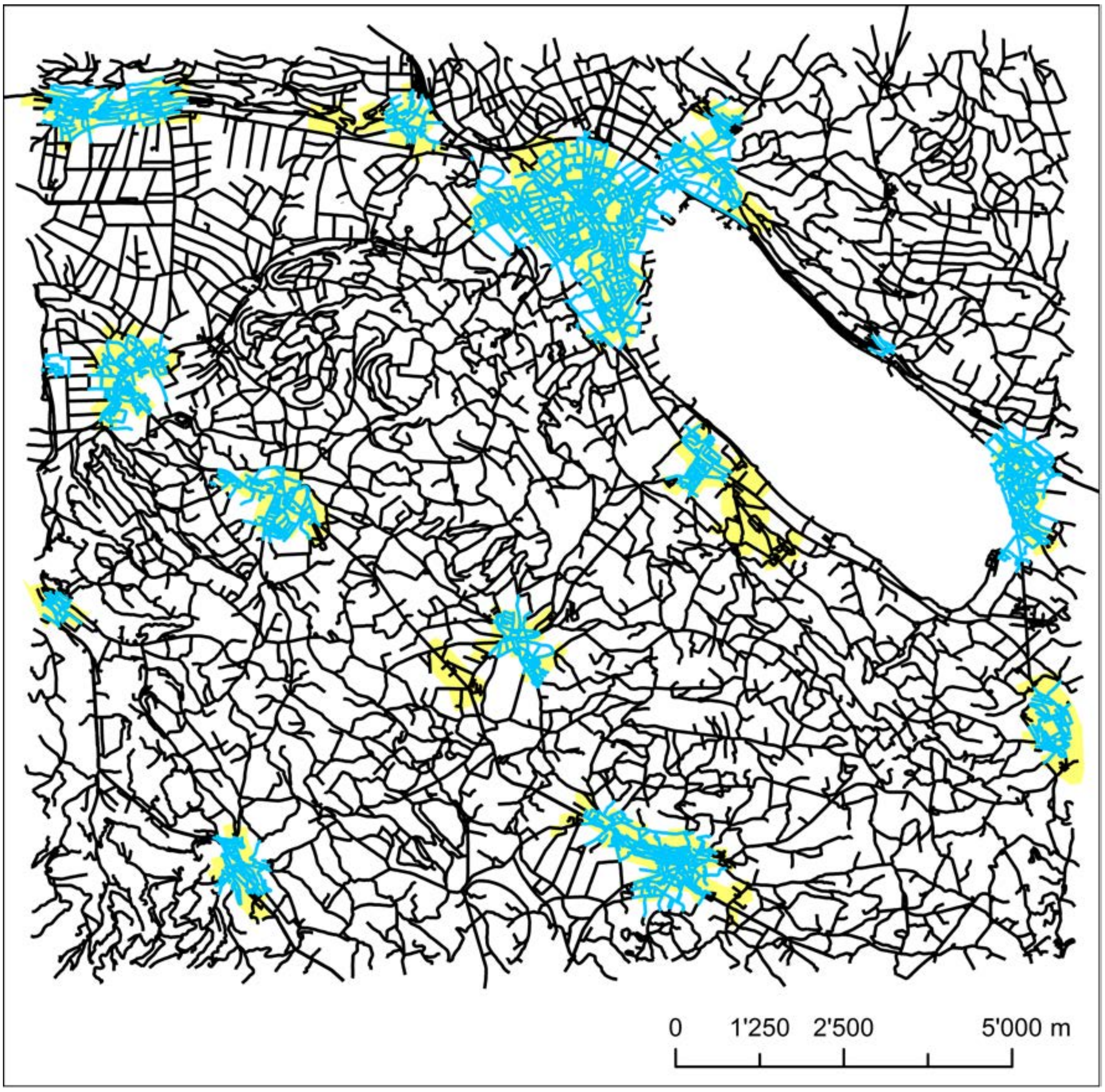

Figure 18. Result of the density algorithm. The blue segments were found by the algorithm whereas the yellow areas model the actual settlement areas (© swisstopo).

As can be seen in Figure 18, the result looks quite promising. Most of the settlement areas were found. More importantly, there are no clusters found by the density algorithm in rural areas (i.e. no false positives). However, there are parts of two small settlement areas that were not discovered by the algorithm. One of these areas could be explained by its location next to a lake, which creates an 'edge effect' situation for the density algorithm. Both areas are small and thus not really essential for the density variation in the resulting road network.

Final results

Table 1 summarizes some properties for the four test areas whereas the properties of the generalization results are shown in Table 2. As can be seen, the selection percentages in all test areas are around 70 %. Constraint H1 was thus met using the iterative mechanism described before. The selection percentages for the total length of all segments are slightly higher, particularly for the EN test area. EN represents part of the Engadine valley, in a very mountainous area on the Swiss Alps. The segments that were deleted are mostly very short segments, thus resulting in a higher selection percentage for the length of the segments.

Furthermore, all segments that must not be deleted because they contain certain important attributes have been selected in the results, hence fulfilling constraint H2. H3 and H4 were also met, due to the logic of the extended stroke-mesh combination algorithm. New dead-end roads only appear because they lead to important POIs, ensured by the accessibility algorithm (thus fulfilling H6). H5 was taken care of using our new roundabout detection algorithm, which was able to reliably spot all the roundabouts.

Evaluating how well constraint S1 is fulfilled is, not surprisingly, rather difficult. The swisstopo experts marked diverse roads in the evaluation that should have been selected but were deleted by the algorithm. In general, however, the experts were very pleased with the result (see also next section for details). Finally, the ratio of the number of segments in urban and rural areas, respectively, could be preserved using the model presented in the previous section, thus fulfilling constraint S2. Only in the test area EN a small difference is noticeable. This is due to the fact that the regression line of our density model does not perfectly fit the data in this case ($R^2 = 0.96$).

Table 1. Description of the four test areas.

Test area	ZU	WT	SU	EN
Number of segments	8,694	30,780	11,129	8,975
Number of linear segments	945	2,951	1,897	1,520
Number of areal segments	7,749	27,829	9,232	7,455
Total length of segments [m]	992,752	4,368,784	1,582,041	2,183,795
Total length of linear segments [m]	119,009	502,478	302,108	429,530
Total length of areal segments [m]	873,743	3,866,306	1,279,933	1,754,265
Number of non-removable segments*	588	1,665	684	445
Number of dead-end roads	801	2,507	1,533	1,302
Ratio of number of segments in urban and rural areas	0.988	0.511	0.517	0.258

* segments that must be selected for the 1:50,000 scale (highways, connection road, etc.).

Table 2. Description of the four generalized test areas.

Test area	ZU	WT	SU	EN
Number of segments	6,092	21,532	7,846	6,323
Selection percentage [%]	70.1	70.0	70.5	70.5
Number of linear segments	278	861	875	833
Number of areal segments	5,814	20,671	6,971	5,490
Total length of segments [m]	718,145	3,132,677	1,153,835	1,705,090
Selection percentage [%]	72.3	71.7	72.9	78.1
Total length of linear segments [m]	58,009	236,991	202,414	351,282
Total length of areal segments [m]	660,136	2,895,686	951,421	1,353,808
Number of non-removable segments*	588	1,665	684	445
Number of dead-end roads	146	474	460	393
Number of *new* dead-end roads	3	15	10	11
Ratio of number of segments in urban and rural areas	0.988	0.511	0.517	0.256

* segments that must be selected for the 1:50,000 scale (highways, connection road, etc.).

Additional results are provided as supplemental online material. Specifically, all of our four test areas are depicted. Furthermore, the results of our algorithm with all the extensions for all of our test areas are shown. Finally, two additional evaluation plots are included that show

how the expert cartographers have evaluated our results (see also next section). Additionally, the full detail about this research can be found in Benz (2013).

Our algorithm was implemented as a prototype in Java. The source code is available under the following URL: https://stbe@bitbucket.org/stbe/selectiontool.git.

Evaluation by swisstopo experts

The results of the algorithm with all the extensions included were evaluated independently by two swisstopo expert cartographers. The two experts have proficient skills and decades of professional experience (38 years and 24 years, respectively) in the manual selection of road networks. Each test area was evaluated separately. The evaluation phase lasted a full day and consisted of two parts. Thus, the small number of evaluators is compensated by their depth of expertise and the great detail and care that was spent on the evaluation (as documented in the sample evaluation plots included in the supplemental online material). On the other hand, it has to be noted that the experts invariably introduce a bias towards the swisstopo style of map making.

The first, shorter part of the expert evaluation consisted of a questionnaire for each of the four test areas. It involved six multiple choice questions that handle qualitative criteria of the road network selection solutions. The experts could rate the different criteria as good, acceptable, bad or useless. Examples of questions are 'How do you rate network pruning in urban areas?' or 'How do you rate the preservation of network structure?'. In general, the answers to the questionnaire established that the solutions are of very good quality. None of the experts rated a criterion as bad or useless. All criteria were rated as either good or acceptable (Benz 2013).

In the second part of the evaluation, the experts were presented the solutions, printed on a plot at the actual target scale of 1:50,000. The experts then examined the selected road network with a special magnifying glass. The goal was to mark and comment on situations that were suboptimal and where the experts would have selected differently. Furthermore, areas that came out especially well were also marked. A small extract of an evaluation plot is depicted in Figure 19 in order to give an impression of the nature and degree of detail of comments made. The roads colored in white are those that were omitted. The retained roads are shown using non-white colors and different degrees of line width, depending on attribute (e.g. highways, connection roads, hiking trails, etc.). Experts marked up the plots using fluorescent markers. Areas that the experts thought were particularly well generalized are circled in green, whereas areas or individual road segments that would have been treated differently by the experts are marked with pink marker, and annotated by some explanatory text. In Figure 19, the annotation of the settlement area circled in pink (lower left) says that this settlement could have been generalized to a greater extent. The annotations of the individual road segments all say that they should have been retained. Once again, we would like to refer the reader to the supplemental online material, where two full evaluation plots (out of eight in total) are provided as examples.

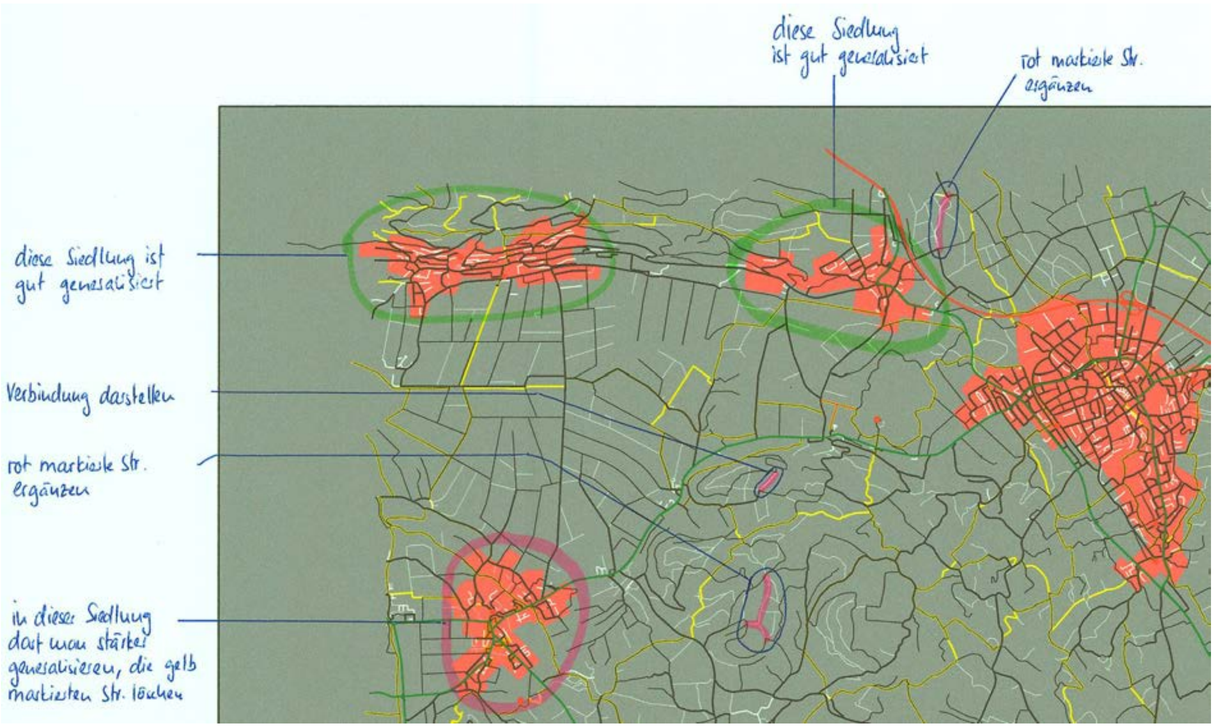

Figure 19. Example of an evaluation plot (© swisstopo).

Generally, the results were rated as very good by the swisstopo experts. In fact, they were positively surprised how well the extended algorithm performed. The experts established that about 5 % to 10 % of the roads would need to be corrected manually, which is, as they remarked, a very low percentage. The results looked so promising that the responsible managers at swisstopo decided to incorporate our software in their map production process.

However, there are also some points that, according to the swisstopo experts, could be improved. For instance, the algorithm selects linear segments (i.e. strokes in linear hierarchies) based on their length or their attributes. If a dead-end road is shorter than the stroke length threshold, the stroke is omitted. However, there are cases, as shown in Figure 20, where there are multiple parallel dead-end roads that were omitted by the algorithm (colored in red) due to their short length. Feedback by the swisstopo experts revealed that in such cases, one or two of those dead-end roads could be selected as representatives in order to show the overall pattern, especially if those roads lead to buildings. Therefore, an extension should be considered to also incorporate a building feature class or to use a method that detects such patterns automatically (e.g. Heinzle and Anders 2007) and handle such dead-end roads differently.

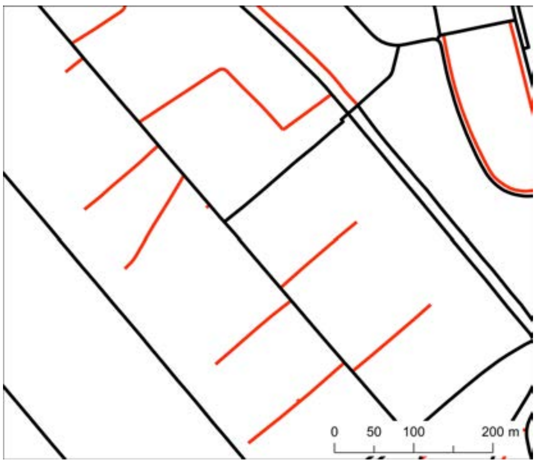

Figure 20. Parallel and omitted dead-end roads (© swisstopo).

Another interesting fact is that most of the time, the experts marked roads which should not have been omitted but selected. The opposite, a road that should have been omitted but was selected by the algorithm, occurred only rarely. Inherently, this could mean that the selection percentage of 70 % (constraint H1) is too low. Further analysis could show whether a higher selection percentage (e.g. 75 %) would select those roads that were marked by swisstopo experts.

Conclusion

This paper reported on a research project (Benz 2013) that was pursued in collaboration with swisstopo, the national mapping agency of Switzerland. The aim was to develop an algorithm for the automated selection of a road network for a target scale of 1:50,000. The road network TLM3D (Topographic Landscape Model 3D), a detailed and dense spatial database with a high spatial resolution in the scale range of 1:10,000, was used as basis for the automated selection. A set of requirements and constraints were defined by expert cartographers at swisstopo; however, they were defined such that they could be adapted to the specific requirements at other mapping agencies, thus offering an opportunity to generalize the methods developed in this work.

Early in the process, three algorithms from the literature were implemented, tested and analyzed on four different test areas. The analysis revealed that the integrated stroke-mesh approach by Li and Zhou (2012) produces the most feasible results. Nevertheless, several difficulties remained and several constraints set by swisstopo experts were violated. Thus, our work sought to obtain better results by introducing five extensions that eliminate the deficiencies of the basic stroke-mesh algorithm. Three of these extensions apply directly to the inner workings of the basic stroke-mesh algorithm. Additionally, two extensions were introduced that allow incorporating external feature classes.

Inter alia, the results were evaluated intensively by swisstopo experts. Their feedback revealed that our extended stroke-mesh algorithm generates very good results and only minimal manual post-processing is necessary. Because of that, swisstopo have decided to incorporate our algorithm in their map production process.

While our extended algorithm was designed to be transferrable and adaptable to the situation at other NMAs, future work will have to establish empirically whether that assumption holds. Furthermore, the applicability for smaller target scales should be evaluated. As a concurrent project (Weiss 2013; Weiss and Weibel 2013) dealing with road network selection for the target scale of 1:200,00 has established, in the scale transition to smaller scales, the attention shifts from local details to the larger network structures and to the overall network topology. Thus, an approach based on a network centrality measure seems more appropriate (Weiss and Weibel 2013). However, where exactly the limits of applicability are between these two approaches will need to be established. Finally, it could also be studied how our extended approach performs if the aim is not to select a certain percentage of segments for the target scale but to use a threshold of overall length of the selected road network.

Acknowledgements

This article represents a largely extended version of a short paper that was originally presented at the 16th ICA Workshop on Generalization and Multiple Representation in Dresden, Germany (Benz and Weibel 2013). We would like to thank swisstopo, in particular Dominik Käuferle, Stefan Wullschleger, Daniel Josi and Marianne Berger, for providing the data and their inputs, as well as the detailed evaluation. Thanks are also due to Roy Weiss for conceptual discussions regarding the programming part of the project.

References

Beard, M. K. 1991. “Constraints on Rule Formation.” In *Generalization: Making Rules for Knowledge Representation*, edited by B.P. Buttenfield and R.B. McMaster, 121-135. London: Longman Group.

Benz, S. 2013. “Algorithmen zur automatischen Selektion von Strassennetzwerken – Lösungen für den Übergang vom Topografischen Landschaftsmodell (TLM) zu Massstab 1:50‘000.” Master Thesis, Dept. of Geography, University of Zurich.

Benz, S.A., and R. Weibel. 2013. “Road Network Selection Using an Extended Stroke-Mesh Combination Algorithm.” Paper presented at the Workshop on the 16th ICA Generalisation Workshop, Dresden, August 23-24.

Chaudhry, O., and W.A. Mackaness. 2005. “Rural and Urban Road Network Generalisation Deriving 1:250,000 from OS MasterMap.” In *Proceedings, 22th International Cartographic Conference*, La Coruña, July 10-16.

Chen, J., Y. Hu, Z. Li, R. Zhao, and L. Meng. 2009. „Selective Omission of Road Features Based on Mesh Density for Automatic Map Generalization.“ *International Journal of Geographical Information Science* 23 (8): 1013-1032.

Edwardes, A.J., and W.A. Mackaness. 2000. “Intelligent Generalisation of Urban Road Networks.” In *Proceedings, GIS Research UK 2000 Conference*, York, April 5-7.

Ester, M., H.-P. Kriegel, J. Sander, and X. Xu. 1996. „A Density-based Algorithm for Discovering Clusters in Large Spatial Databases with Noise.“ In *Proceedings, 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96)*, Portland, OR, August 2-4.

Harrie, L., and R. Weibel. 2007. “Modelling the Overall Process of Generalisation.” In *Generalisation of Geographic Information: Cartographic Modelling and Applications*, edited by W.A. Mackaness, A. Ruas, and L.T. Sarjakoski, 67-88. Oxford: Elsevier.

Hart, P. E., N.J. Nilsson, and B. Raphael. 1968. “A Formal Basis for the Heuristic Determination of Minimum Cost Paths.” *IEEE Transactions on Systems Science and Cybernetics* 4(2): 100–107.

Heinzle, F., K.H. Anders, and M. Sester. 2005. “Graph Based Approaches for Recognition of Patterns and Implicit Information in Road Networks.” In *Proceedings, 22th International Cartographic Conference*, La Coruña, July 10-16.

F. Heinzle and K.-H. Anders. 2007. “Characterising Space via Pattern Recognition Techniques: Identifying Patterns in Road Networks.” In *Generalisation of Geographic Information: Cartographic Modelling and Applications*, edited by W.A. Mackaness, A. Ruas, and L.T. Sarjakoski, 233-253. Oxford: Elsevier.

Jiang, B. 2009. “Street Hierarchies: A Minority of Streets Account for a Majority of Traffic Flow.” *International Journal of Geographical Information Science* 23 (8): 1033-1048.

Jiang, B., and C. Claramunt. 2004. "A Structural Approach to the Model Generalization of Urban Street Network." *GeoInformatica* 8 (2): 157-171.

Jiang, B., and L. Harrie. 2004. "Selection of Streets from a Network Using Self-Organizing Maps." *Transactions in GIS* 8 (3): 335-350.

Li, Z., and Q. Zhou. 2012. "Integration of Linear and Areal Hierarchies for Continuous Multi-Scale Representation of Road Networks." *International Journal of Geographical Information Science* 26 (5): 855-880.

Liu, X., F.B. Zhan, and T. Ai. 2010. "Road Selection Based on Voronoi Diagrams and "Strokes" in Map Generalization." *International Journal of Applied Earth Observation and Geoinformation* 12 (2): 194-202.

Mackaness, W.A. 1995. "Analysis of Urban Road Networks to Support Cartographic Generalization." *Cartography and Geographic Information Systems* 22(4): 306-316.

Mackaness, W.A., and G.A. Mackechnie. 1999. "Automating the Detection and Simplification of Junctions in Road Networks." *GeoInformatica* 3 (2): 185-200.

Mackaness, W.A., and M.K. Beard. 1993. "Use of Graph Theory to Support Generalisation." *Cartography and Geographic Information Systems* 20 (4): 210-221.

Morisset, B., and A. Ruas. 1997. "Simulation and Agent Modelling for Road Selection in Generalisation." In *Proceedings, 18th International Cartographic Conference*, Stockholm, June 23-27.

Porta, S., P. Crucitti, and V. Latora. 2006. "The Network Analysis of Urban Streets: A Dual Approach." *Physica A: Statistical Mechanics and its Applications* 369 (2): 853-866.

Richardson, D.E., and R.C. Thomson. 1996. "Integrating Thematic, Geometric and Topologic Information in the Generalization of Road Networks." *Cartographica* 33 (1): 75–83

Thom, S. 2005. "A Strategy for Collapsing OS Integrated Transport Network(tm) Dual Carriageways." Paper presented at the ICA Workshop on Generalisation and Multiple Representation, La Coruña, July 7-8.

Thomson, R.C., and D.E. Richardson. 1999. "The Good Continuation Principle of Perceptual Organisation Applied to the Generalisation of Road Networks." In *Proceedings, 19th International Cartographic Conference*, Ottawa, August 14-21.

Thomson, R.C., and R. Brooks. 2000. "Efficient Generalisation and Abstraction of Network Data Using Perceptual Grouping." *http://www.geocomputation.org/2000/GC029/Gc029.htm*.

Tomko, M., S. Winter, and C. Claramunt. 2008. "Experiential Hierarchies of Streets." *Computers, Environment and Urban Systems* 32 (1): 41-52.

Touya, G. 2010. “A Road Network Selection Process Based on Data Enrichment and Structure Detection.” *Transactions in GIS* 14 (5): 595-614.

Walter, V. 2008. “Automatic Interpretation of Vector Databases with a Raster-Based Algorithm.” In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVII, Part B2, Commission 2*, Beijing, July 3-11.

Weibel, R., and G.H. Dutton. 1998. “Constraint-Based Automated Map Generalization.” In *Proceedings, 8th International Symposium on Spatial Data Handling*, Vancouver, July 11-15.

Weiss, R. 2013. “Road Network Selection for Small-Scale Maps Using an Improved Centrality Approach.” Master Thesis, Dept. of Geography, University of Zurich.

Weiss, R., and R. Weibel. 2013. “Road Network Selection for Small-Scale Maps Using an Improved Centrality Approach.” Paper presented at the Workshop on the 16th ICA Generalisation Workshop, Dresden, August 23-24.

Yang, B., X. Luan, and Q. Li. 2011. “Generating Hierarchical Strokes from Urban Street Networks Based on Spatial Pattern Recognition.” *International Journal of Geographical Information Science* 25 (12): 2025-2050.

Zhang, Q. 2004. “Modeling Structure and Patterns in Road Network Generalization.” Paper presented at the Workshop on Generalisation and Multiple Representation, Leicester, August 20-21.